PCT WELTORGANISATION FOR GHISTIGES EIGENTUM
Internationales Boro
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 6: (11) Internationale Veröffentlichungmummer: WO 98/10352 A1 G06F 15/80 (43) Internationales

(21) Internationales Aktenzelchen:

PCT/DE97/01497

(22) Internationales Anmeldedatum:

15. Juli 1997 (15.07.97)

(81) Bestimmungsstaaten: JP, US, europäisches Patent (AT, BE, CH, DR, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL,

12. Marz 1998 (12.03.98)

(30) Prioritätsdaten:

196 35 758.6

3. September 1996 (03.09.96)

DE

Veröffentlicht

Veröffentlichungsdatum:

Mit internationalem Recherchenbericht.

(71) Anmelder (für alle Bestimmungsstaaten ausser US): SIEMENS AKTIENGESELLSCHAPT [DE/DE]; Wittelsbacherplatz 2, D-80333 München (DE).

(72) Erfinder; und

(75) Brinder/Ammelder (nur für US): NEUNEIER, Ralf [DE/DE]; Utzschneiderstrasse 10, D-80469 München (DB). ZIMMER-MANN, Hans-Georg [DE/DE]; Schiffbauerweg 6 A, D-82319 Starnberg (DE).

(54) Title: METHOD AND DEVICE FOR COMPUTER ASSISTED GENERATION OF AT LEAST ONE ARTIFICIAL TRAINING DATA VECTOR FOR A NEURONAL NETWORK

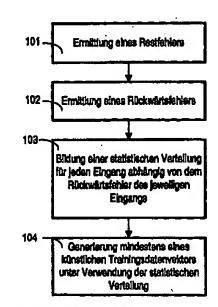
(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUR RECHNERGESTÜTZTEN GENERIERUNG MINDESTENS EINES KÜNSTLICHEN TRAININGSDATENVEKTORS FÜR EIN NEURONALES NETZ

(57) Abstract

A residual error (101) is determined after neuronal network training has occurred. A backward error (102) is determined from the residual error. Artificial training data vectors are generated from a statistic random process with an underlying statistic distribution, wherein the respective backward error is considered for a neuronal network input (103, 104).

(57) Zusummenfassung

Nach erfolgtem Training eines neuronalen Netzes wird ein Restfehler ermittelt (101). Aus dem Restfehler wird ein Rückwärtsfehler ermittelt (102). Künstliche Trainingsdatenvektoren werden aus einem statistischen Zufallsprozeß generiert, dem eine statistische Verteilung zugrundeliegt, bei der der jeweilige Rückwärtsfehler für einen Eingung des neuronalen Netzes berücksichtigt wird (103, 104).



101 — DETERMINATION OF AN UNDETECTED ERROR
102 — DETEMBRATION OF A BACKWARD ERROR
103 — FORMATION OF A STATISTIC DISTRIBUTION FOR EACH INPUT
DEPENDENT ON THE BACKWARD ERROR OF THE RESPECTIVE INPUT
104 — GENERATION OF AT LEAST ONE ARTUFICIAL TRAINING DATA VECTOR
USING STATISTIC DISTRIBUTION

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

	AL	Albanien	¥8	Spanien	LS	Lexotho	81	Slowenien
	AM	America	n	Finaland	LT	Litruen	SK	Slowakei
	AT	Osterreich	FR	Prankreich	LU	Loxemburg	8N	Senegal
ŀ	AU	Anstrallen	GA	Gabun	LV	Lettland	8 Z	Swasiland
	AZ	Aserbaidachan	GB	Vereinigtes Königreich	MC	Monaco	TD	Techad
	BA	Bornier-Hemogowina	GE	Georgien	MD	Republik Moldau	TG	Togo
	BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tedechikistan
	BE	Belgion	GN	Guinea	MK	Die chemalige jugoslawische	TM	Turkmenistan
	BF	Burkina Faso	GR	Oriochealand		Republik Mazedonien	TR	Türkel
	BG	Bulgarien	HU	Umpen	ML	Mali	TT	Trinidad und Tobago
l	BJ	Benin	IE	Irland	MN	Mangolei	UA	Ukraine
ı	BR	Brasilien	IL.	Israel	MR	Mauretanien	UG	Uganda
ı	BY	Belanu	18	la land	MW	Malawi	US	Versinisto Staaten von
ı	CA	Kanada	iT	Italien	MX	Mexiko		Amerika
ı	Œ.	Zentralafrikanische Republik	115	Japan	NB	Niger	UZ	Usbekistan
l	œ	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
		Schweiz	KG	Kirgisisan	NO	Norwegen	YU	Jugoslawion
ı	CH	Câte d'Ivoire	KP	Demokratische Volksrepublik	NZ	Neusesland	ZW	Zimbabwe
l			202	Korta	PL	Polen		
١	CM	Kameran	KR	Republik Korza	PT	Portugal		
l	CN	China	KZ	Kasachstan	RO	Ruminien		
١	CU	Kuba				Russische Röderstion		
ı	CZ	Tachechische Republik	LC	St. Locia	RU			
l	DB	Deutschland	Ц	Liechtemtein	SD	Sudan		
ı	DK	Dinemark	LK	Szi Lanka	SE	Schweden		
l	22	Dulland	1.0	1.lberia	5G	Singapor		

Beschreibung

Verfahren und Vorrichtung zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein neuronales Netz

1. Technischer Hintergrund

10

15

Neuronale Netze lernen mit Hilfe von Trainingsdaten. Die Trainingsdaten sind in vielen Anwendungsgebieten sehr verrauscht, beispielsweise bei der Modellierung von Finanzdaten wie Aktien- oder Währungskursen. Somit erhalten die Trainingsdaten zufällige Störungen, die nichts mit der eigentlich zu modellierenden Systemdynamik zu tun haben.

Durch die Approximationsfähigkeit der neuronalen Netze kann aber auch die transiente Struktur des zufälligen Rauschens gelernt werden. Dieses Phänomen wird als Übertrainieren des neuronalen Netzes bezeichnet. Durch ein übertrainiertes neuronales Netz wird in sehr verrauschten Systemen der Lernvorgang des neuronalen Netzes erheblich behindert, da die Verallgemeinerungsfähigkeit des neuronalen Netzes negativ beeinflußt wird.

Dieses Problem gewinnt in Anwendungsgebieten an Bedeutung, in denen nur eine geringe Anzahl von Trainingsdatenvektoren zur Adaption des neuronalen Netzes an die Anwendung, d.h. die zu modellierende, durch die Trainingsdatenvektoren repräsentierte Funktion, verfügbar ist.

Vor allen in diesen Anwendungsgebieten, aber auch allgemein in einem Trainingsverfahren eines neuronalen Netzes ist es vorteilhaft, zusätzliche Trainingsdatenvektoren künstlich zu

2

generieren, um somit eine größere Trainingsdatenmenge zu erhalten.

2. Stand der Technik

Es ist bekannt, die Generierung der künstlichen Trainingsdatenvektoren durch Verrauschen der verfügbaren Trainingsdatenvektoren der Trainingsdatenmenge durchzuführen. Es ist in diesem Zusammenhang aus dem Dokument [1] bekannt, die Trainingsdatenmenge mit gaußförmigem Rauschen mit dem Mittelwert 0 und einer Varianz G, die für alle Eingänge des neuronalen Netzes auf den gleichen Wert gesetzt wird, zu ermitteln.

Aus [4] ist es bekannt, Trainingsdaten durch Einführung zusätzlichen Rauschens zu generieren. Dabei ist es bekannt, die sogenannte Jackknife-Prozedur einzusetzen.

Dieses Verfahren weist jedoch einige Nachteile auf.

Dadurch, daß zur Generierung der zusätzlichen Trainingsdaten-20 vektoren als statistische Verteilung, die zur Generierung verwendet wird, ein gaußförmiges Rauschen mit einer Varianz verwendet wird, die für alle Eingänge des neuronalen Netzes auf den gleichen Wert gesetzt wird, werden Trainingsdatenvektoren neu generiert, die keinerlei Aussage über das zu model-25 lierende System enthalten. Die Trainingsdatenvektoren enthalten ferner keinerlei Information über das tatsächliche, dem System zugrundeliegende Rauschen. Somit wird zwar die Trainingsdatenmenge vergrößert, diese muß aber nicht den Lernvorgang unterstützen, da ein fest vorgegebenes Rauschen, das mit der eigentlichen Systemdynamik nichts zu tun hat, zum Trai-30 nieren des neuronalen Netzes verwendet wird. Damit kann es dann trotzdem zum Übertrainieren kommen.

Grundlagen über neuronale Netze sind beispielsweise aus dem 35 Dokument [2] bekannt.

3

Grundlagen über Verwendung der neuronalen Netze in der Ökonomie sind beispielsweise aus dem Dokument [3] bekannt.

3. Kurzbeschreibung der Erfindung

5

Somit liegt der Erfindung das Problem zugrunde; künstlich neue Trainingsdatenvektoren für ein neuronales Netz zu bilden, wobei jedoch ein Übertrainieren des neuronalen Netzes vermieden wird.

10

Das Problem wird durch das Verfahren gemäß Patentanspruch 1 sowie durch die Vorrichtung gemäß Patentanspruch 9 gelöst.

Bei dem Verfahren wird nach dem Training des neuronalen Netzes mit verfügbaren Trainingsdatenvektoren einer Trainingsda-15 tenmenge ein Restfehler ermittelt. Aus dem Restfehler wird beispielsweise unter Verwendung eines Gradientenabstiegs-Verfahren ein eingangsbezogener Rückwärtsfehler ermittelt. Die Ermittlung des Rückwärtsfehlers entspricht der üblichen Vorgehensweise während des Trainings eines neuronalen Netzes 20 zur Adaption der einzelnen Gewichte des neuronalen Netzes. Ist der eingangsbezogene Rückwärtsfehler ermittelt worden, wird unter Berücksichtigung des jeweiligen Rückwärtsfehlers eine dem jeweiligen Eingang zugeordnete statistische Verteilung generiert, und der künstliche Trainingsdatenvektor wird 25 unter Berücksichtigung der jeweiligen statistischen Verteilung an den Eingängen des neuronalen Netzes erzeugt.

Mit diesem Verfahren ist es möglich, zusätzliche Trainingsdatenvektoren zu generieren, die eine Information über das neuronale Netz und die Struktur des neuronalen Netzes aktuell nach dem Training des neuronalen Netzes mit den verfügbaren Traindingsdatenvektoren enthalten.

Dies führt dazu, daß die künstlich erzeugten Trainingsdatenvektoren von dem Rückwärtsfehler, der nach dem Training des neuronalen Netzes noch existiert, abhängig sind und somit von

4

der Performanz des neuronalen Netzes abhängen. Dies führt dazu, daß die erzeugten Trainingsdatenvektoren Information über das zu modellierende System enthalten. Auf diese Weise kann ein Übertrainieren des neuronalen Netzes durch die zusätzlichen Trainingsdatenvektoren vermieden werden.

Die Vorrichtung gemäß Patentanspruch 9 weist eine Recheneinheit auf, die derart eingerichtet ist, daß die oben beschriebenen Verfahrensschritte durchgeführt werden.

10

Vorteilhafte Weiterbildungen der Erfindung ergeben sich aus den abhängigen Ansprüchen.

Es ist vorteilhaft, bei einer Online-Approximation des neuronalen Netzes, die auch als Online-Training bezeichnet wird,
auch die jeweilige statistische Verteilung dem veränderten
Trainingsdatensatz anzupassen. Dadurch wird das zu modellierende System noch genauer durch das neuronale Netz modelliert.

20

4. Kurze Figurenbeschreibung

In den Figuren ist ein Ausführungsbeispiel der Erfindung dargestellt, welches im weiteren näher erläutert wird.

25

Es zeigen

Fig. 1 ein Ablaufdiagramm, in dem die einzelnen Verfahrensschritte des Verfahrens dargestellt sind;

30

Fig. 2 eine Skizze, in der eine Rechneranordnung dargestellt ist, mit der das Verfahren durchgeführt werden kann.

5. Figurenbeschreibung

In Fig. 1 sind die einzelnen Verfahrensschritte des Verfahrens dargestellt.

In einem ersten Schritt (101) wird nach abgeschlossenem Training des neuronalen Netzes NN mit einer beliebigen Anzahl von Trainingsdatenvektoren TDVi einer Trainingsdatenmenge TDM von einem Rechner R ein Restfehler RE ermittelt.

Der Restfehler RE wird beispielsweise bei m Trainingsdatenvektoren TDV mit den Trainingsdatenvektoren TDV zugeordneten
Sollwerten ti, wobei mit einem Index i jeweils der Trainingsdatenvektor TDVi eindeutig gekennzeichnet wird, auf folgende
Weise ermittelt:

15 RE =
$$\sum_{i=1}^{m} (t_i - y_i)^2$$
 (1).

Es kann jedoch selbstverständliche jede beliebige andere Vorschrift zur Bildung des Restfehlers RE nach dem Training des neuronalen Netzes NN verwendet werden.

Nach Ermittlung des Restfehlers wird ein Rückwärtsfehler RFj ermittelt (102). Die Bestimmung des Rückwärtsfehlers RF kann beispielsweise unter Verwendung eines Gradientenabstiegs-Verfahrens nach den Eingangssignalen der neuronalen Netze NN erfolgen.

Im weiteren wird folgende Nomenklatur für den Trainingsdatenvektor TDVi, sowie die Eingangsdaten des neuronalen Netzes NN, die sich aus den Trainingsdatenvektoren TDVi und einem Rauschtermvektor $\underline{\mathbf{e}}$ ergeben, verwendet:

Trainingsdatenvektor TDVi: $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n)$

Rauschtermvektor: $\underline{e} = (e_1, \ldots, e_j, \ldots, e_n)$

. 20

25

6

Eingangsdatenvektor: $\underline{x} = (x_1+e_1, \dots, x_j+e_j, \dots, x_n+e_n)$

Mit n wird eine Anzahl von Komponenten der Trainingsdatenvektoren TDVi, der Rauschtermvektoren <u>e</u> bzw. der Eingangsdatenvektoren <u>z</u> bezeichnet.

Der Rückwärtsfehler RFj wird für jeden Eingang j des neuronalen Netzes NN individuell ermittelt, wobei mit einem Index j jeweils eine Komponente des Eingangsdatenvektors z bzw. ein Eingang des neuronalen Netzes NN eindeutig bezeichnet wird,.

Dies erfolgt beispielsweise nach dem bekannten Gradientenabstiegs-Verfahren zum Trainieren des neuronalen Netzes NN. Somit ergibt sich der Rückwärtsfehler RF des Eingangsdatenvektors z aus den jeweiligen partiellen Ableitungen des Restfehlers RE nach den einzelnen Eingangssignalen zj.

$$RFj = \frac{\partial E^{t}}{\partial z_{j}^{t}}$$
 (2).

20 Mit dem Symbol t wird jeweils ein Zeitpunkt eindeutig gekennzeichnet, zu dem die Ermittlung des Rückwärtsfehlers RFj erfolgt.

Unter Berücksichtigung des Rückwärtsfehlers RFj wird nun in einem dritten Schritt (103) eine statistische Verteilung S_j für jeden Eingang j des neuronalen Netzes NN gebildet.

Als statistische Verteilung $S_{\hat{j}}$ kann allgemein jede beliebige statistische Verteilung zur Beschreibung eines Zufallsprozesses verwendet werden.

Im folgenden werden zur einfacheren Darstellung jedoch nur eine uniforme statistische Verteilung und eine gaußförmige statistische Verteilung $S_{\bar{j}}$ näher erläutert.

30

Für die uniforme Verteilung ergibt sich eine uniforme Verteilungsbreite s_i^t beispielsweise nach folgender Vorschrift:

$$s_{j}^{t} = \frac{1}{t} \sum_{k=1}^{t} \left| \frac{\partial E^{k}}{\partial z_{j}^{k}} \right|$$
 (3).

Der Rauschterm ej^t, der zur Bildung mindestens eines künstlichen Trainingsdatenvektors KTDV verwendet wird, liegt im Bereich des Intervalls:

10
$$e_j^t \in \left[-s_j^t, s_j^t\right]$$
 (4).

Bei einer gaußförmigen Verteilung ergibt sich die Rauschbreite \mathbf{s}_i^t nach der folgenden Vorschrift:

15
$$s_j^t = \frac{1}{t} \sum_{k=1}^t \left(\frac{\partial E^k}{\partial z_j^k} \right)^2$$
 (5).

Der Rauschterm ej^t ergibt sich für diesen Beispielsfall nach folgender Vorschrift:

$$e_{j}^{t} \in N\left(0, \sqrt{s_{j}^{t}}\right) \tag{6}$$

Mit $N(0, \sqrt{s_j^t})$ wird ein normal verteilter Zufallsprozeß mit einem Mittelwert 0 und der Standardabweichung $\sqrt{s_j^t}$ bezeichnet.

Unter Verwendung der jeweiligen statistischen Verteilung $S_{\rm j}$ wird der mindestens eine künstliche Trainingsdatenvektor KTDV generiert (104).

Anschaulich bedeutet dies, daß der künstliche Trainingsdatenvektor KTDV durch den jeweils durch die statistische Verteilung S_j beschriebenen Zufallsprozeß generiert wird.

5 Durch diese Vorgehensweise haben die künstlichen Trainingsdatenvektoren KTDV statistisch die gleichen Eigenschaften wie die original verfügbaren Trainingsdatenvektoren TDVi.

Dies kann durch Verrauschen mit einem einheitlichen
Rauschprozeß, wie dieser aus dem Dokument [1] bekannt ist,
nicht erreicht werden.

Anschaulich läßt sich das Verfahren folgendermaßen beschreiben. Die Eingangsdaten, repräsentiert durch die original verfügbaren Trainingsdatenvektoren TDVi, werden mit einem Zufallsprozeß verrauscht, der für jeden Eingang abhängig ist von der jeweiligen Gradienten-Fehlerfunktion nach den Eingangssignalen.

Die jeweilige Rauschbreite \mathbf{s}_{j}^{t} kann anschaulich folgendermaßen interpretiert werden:

- Kleine Werte von $\mathbf{s}_{\mathbf{j}}^{\mathsf{t}}$ werden durch Eingänge erzeugt, die eine gute Abbildung auf die Trainingsdaten erzeugen.
- Kleine Werte der Rauschbreite $\mathbf{s}_{\mathbf{j}}^{\mathbf{t}}$ können aber auch auf unwichtige, für das Training des neuronalen Netzes NN nicht relevante Eingänge hinweisen.
- 30 In beiden Fällen ist es im Sinne der Aufgabe, daß kein oder nur wenig Rauschen zu den ursprünglich vorhandenen Trainingsdatenvektoren TDVi hinzugefügt wird.
- Große Werte der Rauschbreite s_j^t deuten darauf hin, daß der 35 Eingang j erhebliche Bedeutung aufweist, aber das neuronale

9

Netz NN bisher nicht in der Lage ist, eine adāquate Abbildung zu lernen.

- Rauschen auf Eingängen mit einer großen Rauschbreite st 5 "versteift" das neuronale Netz NN, wodurch ein Übertrainieren verhindert wird, was zu einer besseren Verallgemeinerungsfähigkeit des neuronalen Netzes NN führt.

Dieses Verfahren kann besonders vorteilhaft in Anwendungssituationen verwendet werden, in denen nur eine relativ geringe Anzahl von Trainingsdatenvektoren TDVi zum Trainieren des neuronalen Netzes NN verfügbar ist.

Durch das Verfahren kann die Trainingsdatenmenge TDM künstlich erheblich erweitert werden zu einer erweiterten Trainingsdatenmenge KTDM, ohne die eigentliche Systemdynamik zu
verfälschen, da die statistischen Eigenschaften der Trainingsdatenmenge TDM auch in den künstlichen Trainingsdatenvektoren KTDV enthalten sind.

20

Eine typische Anwendungssituation des Verfahrens liegt in der Analyse von Finanzmärkten, beispielsweise von Aktienmärkten oder auch Wertpapiermärkten.

Wird zum Trainieren des neuronalen Netzes NN ein sog. OnlineLernverfahren eingesetzt, bei dem während der Generalisierungsphase des neuronalen Netzes NN weitere Trainingsdatenvektoren TDV ermittelt werden, mit denen das neuronale Netz
NN in der Generalisierungsphase weiter trainiert wird, so ist
es in einer Weiterbildung des Verfahrens vorgesehen, abhängig
von der Änderung der Trainingsdatenmenge TDM, die durch die
Trainingsdatenvektoren TDVi gebildet wird, auch die statistischen Verteilungen S_j anzupassen. Dies kann beispielsweise
für eine uniforme Verteilung nach folgender Vorschrift erfolgen:

10

$$\mathbf{s}_{\mathbf{j}}^{\mathsf{t+1}} = (1 - \alpha) \cdot \mathbf{s}_{\mathbf{j}}^{\mathsf{t}} + \alpha \cdot \left| \frac{\partial \mathbf{E}^{\mathbf{k}}}{\partial \mathbf{z}_{\mathbf{j}}^{\mathsf{k}}} \right| \tag{7}.$$

Für eine gaußförmige statistische Verteilung Sj kann beispielsweise folgende Adaptionsvorschrift eingesetzt werden:

5

$$\left(\mathbf{s}_{j}^{t+1}\right)^{2} = \left(1 - \alpha\right) \cdot \left(\mathbf{s}_{j}^{t}\right)^{2} + \alpha \cdot \left(\frac{\partial \mathbf{E}^{k}}{\partial \mathbf{z}_{j}^{k}}\right)^{2} \tag{8}$$

Mit α wird ein Adaptionsfaktor bezeichnet.

10 Es hat sich als vorteilhaft herausgestellt, als Wert für den Adaptionsfaktor α den Kehrwert der Anzahl der Trainingsdatenvektoren TDVi zu verwenden.

Im Rahmen dieses Dokumentes wurden folgende Veröffentlichungen zitiert:

- [1] C. M. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, U. K., ISBN 0-19-853864-2, S. 346 349, 1994
- [2] G. Deco und D: Obradovic, An Information-Theoretic
 Approach to Neural Computing, Springer Verlag, New York,
 USA, ISBN 0-387-94666-7, S. 23 37, 1996
 - [3] H. Zimmermann und Rehkugler, Neuronale Netze in der Ökonomie, Kapitel 1, Neuronale Netze als Entscheidungskalkül, Vahlen Verlag, München, ISBN 3-8006-1871-0, S. 3 - 20, 1994
 - [4] L. Holmström und P. Koistinen, Using Additive Noise in Back-Propagation Training, IEEE Transactions on Neural Networks, Vol. 3, No. 1, S. 24 - 38, Januar 1992

Patentansprüche

- Verfahren zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein Neuronales Netz
 (NN),
 - bei dem ein Restfehler (RE) nach einem Training des Neuronalen Netzes (NN) mit Trainingsdatenvektoren (TDVi) einer Trainingsdatenmenge (TDM) ermittelt wird (101),
- bei dem aus dem Restfehler (RE) für mindestens einen Eingang (j) des Neuronalen Netzes (NN) ein Rückwärtsfehler (RFj)
 ermittelt wird (102),
 - bei dem jeweils eine dem Eingang (j) zugeordnete statistische Verteilung (S_j) generiert wird, die von der Größe des jeweiligen Rückwärtsfehlers (RFj) abhängig ist (103),
- 15 bei dem der künstliche Trainingsdatenvektor generiert wird unter jeweiliger Verwendung der einem Eingang (j) zugeordneten statistischen Verteilung (S_j) (104).
- Verfahren nach Anspruch 1,
 bei dem der Rückwärtsfehler (RFj) mit einem Gradientenabstiegsverfahren ermittelt wird.
- Verfahren nach Anspruch 2,
 bei dem das Gradientenabstiegsverfahren mit einem Backpropa gation-Verfahren realisiert wird.
 - 4. Verfahren nach einem der Ansprüche 1 bis 3, bei dem die statistische Verteilung (S_j) durch eine uniforme Verteilung gebildet wird, wobei eine Verteilungsbreite der uniformen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.
 - 5. Verfahren nach einem der Ansprüche 1 bis 3, bei dem die statistische Verteilung (S_j) durch eine gaußförmige Verteilung gebildet wird, wobei eine Verteilungsbreite der gaußförmigen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

- 6. Verfahren nach Anspruch 4 oder 5, bei dem bei einer Online-Approximation des Neuronalen Netzes (NN) nach einer beliebigen Anzahl neuer Trainingsdatenvekto-5 ren (TDVi) die statistische Verteilung (S_j) an einen neu ermittelten Rückwärtsfehler angepaßt wird.
- 7. Verfahren nach einem der Ansprüche 1 bis 6 eingesetzt zur Modellierung einer nichtlinearen Funktion, die mit einer geringen Anzahl von Trainingsdatenvektoren beschrieben wird. 10
 - 8. Verfahren nach einem der Ansprüche 1 bis 6 eingesetzt in der Analyse eines Finanzmarktes.
- 9. Vorrichtung zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein Neuronales Netz (NN),
 - mit einer Recheneinheit, die derart eingerichtet ist, daß - ein Restfehler (RE) nach einem Training des Neuronalen Net-
- 20 zes (NN) mit Trainingsdatenvektoren (TDVi) einer Trainingsda-
- tenmenge (TDM) ermittelt wird (101),
 - aus dem Restfehler (RE) für mindestens einen Eingang (j) des Neuronalen Netzes (NN) ein Rückwärtsfehler (RFj) ermittelt wird (102),
- jeweils eine dem Eingang (j) zugeordnete statistische Ver-25 teilung (Sj) generiert wird, die von der Größe des jeweiligen Rückwärtsfehlers (RFj) abhängig ist (103),
- der künstliche Trainingsdatenvektor generiert wird unter jeweiliger Verwendung der einem Eingang (j) zugeordneten statistischen Verteilung (Sj) (104). 30
- 10. Vorrichtung nach Anspruch 9, bei der die Recheneinheit derart eingerichtet ist, daß der Rückwärtsfehler (RFj) mit einem Gradientenabstiegsverfahren ermittelt wird. 35
 - 11. Vorrichtung nach Anspruch 9 oder 10,

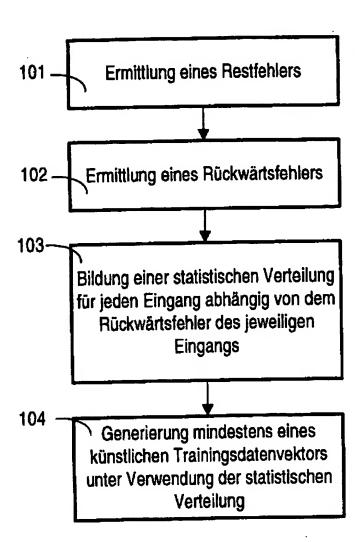
14

bei der die Recheneinheit derart eingerichtet ist, daß die statistische Verteilung (S_j) durch eine uniforme Verteilung gebildet wird, wobei eine Verteilungsbreite der uniformen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

5

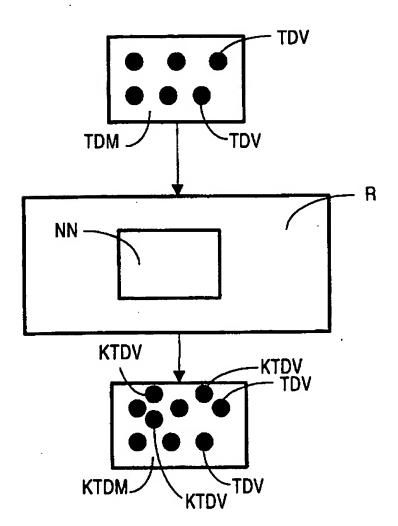
- 12. Vorrichtung nach einem der Ansprüche 9 bis 11, bei der die Recheneinheit derart eingerichtet ist, daß die statistische Verteilung (S_j) durch eine gaußförmige Verteilung gebildet wird, wobei eine Verteilungsbreite der gaußförmigen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.
- 13. Vorrichtung nach Anspruch 11 oder 12, bei der die Recheneinheit derart eingerichtet ist, daß bei einer Online-Approximation des Neuronalen Netzes (NN) nach einer beliebigen Anzahl neuer Trainingsdatenvektoren (TDVi) die statistische Verteilung (Sj) an einen neu ermittelten Rückwärtsfehler angepaßt wird.
- 14. Vorrichtung nach einem der Ansprüche 9 bis 13 eingesetzt 20 zur Modellierung einer nichtlinearen Funktion, die mit einer geringen Anzahl von Trainingsdatenvektoren beschrieben wird.
 - 15. Vorrichtung nach einem der Ansprüche 9 bis 13 eingesetzt in der Analyse eines Finanzmarktes.

FIG 1



2/2

FIG 2



INTERNATIONAL SEARCH REPORT

In. .estional Application No PCT/DE 97/01497

		1	, , .	
A. CLASSI IPC 6	FICATION OF SUBJECT MATTER G06F15/80		-	
According to	o International Patent Classification(IPC) or to both national classifi	cation and IPC		
8. FIELDS	SEARCHED			
Minimum do IPC 6	ocumentation searched (classification system followed by classifica GOSF	tion symbols)		
Occumenta	Intermitted Potent Classification (RPC) or to both national classification and IPC BEARCHD manufacture meached (dasafication system followed by classification symbols) GOSF on searched other than minimum/documentation to the extent that such documents are included in the fields searched at base consulted during the international search (name of data base and, where practical, search terms used) WTS CONSIDERED TO BE RELEVANT Citation of decument, with indication, where appropriate, of the relevant passages WD 95 11486 A (MILES INC) 27 April 1995 See page 10, 1 line 1 – 1 line 32; figure 3 US 5 359 699 A (TONG DAVID W ET AL) 25 October 1994 See Column 4, 11ne 5 – 11ne 42 -/ WP Priest family members are listed in arresz. The last documents are listed in the continuation of box C. X Priest family members are listed in arresz. The last documents are listed in the continuation of box C. X Priest family members are listed in arresz. The last document published after the international filing data or priority data and not in conflex with the exploation but displaying the general state of the safe the international priority data and not in conflex with the exploation but caused to be of particular relevance, the claimed to relevance the claimed haverton or coloring special reason (see specified) The document of pericular relevance the claimed haverton or coloring special reason (see specified) The document of pericular relevance the claimed haverton or coloring special reason (see specified) The document of pericular relevance the claimed haverton or coloring specified pericular relevance the claimed haverton or coloring specified to pericular relevance the claimed haverton or coloring special pericular relevance the claimed haverton or coloring special pericular relevance the claimed haverton or coloring special pericular relevance the claimed to considered to remove an invertible day when the document is base alternational three points and the pericular deciments or the same perior document is b			
Electronic d	tata base consulled during the international search (name of data b	pase and, where practical,	search (erms used)	
C. DOCUM	ENTS CONSIDERED TO BE RELEVANT			
Category *	Citation of document, with indication, where appropriate, of the ri	slevent passages		Relevant to claim No.
A				1,9
A	October 1994	T AL) 25		1,9
X Fun	ther documents are listed in the continuation of box C.	X Patent family	members are listed in	annex.
"A" docum consider "E" earlier filling of "L" docum which citatio "O" docum other "P" docum later t	atagories of clad documents: pertia defining the general state of the art which is not great to be of particular relevance of cocument but published on a raiser the international data ent which may throw doubts on priority claim(s) or is citied to establish the publication date of another in or other special reason (as specified) ent referring to an oral disclosure, use, exhibition or means every published prior to the international filling date but their the priority date claimed.	or priority date an cited to understan invention "X" document of partic carnot be consist involve an invention "Y" document of partic carnot be consist document is comment is comment in the art. "&" document member	Ind not in conflict with the distribution of the claim of	e application but ry underlying the imed invention a considered to ment is taken alone imed invention other step when the a other such docu- to a person stided mily
	B December 1997	16/12/1		
Name and	mailing address of the ISA European Patent Office, P.B. 5616 Patentiesn 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo ni, Fex: (-31-70) 340-3016	Authorized officer Schenke	els, P	

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

PCT/DE 97/01497

PCT/DE 97/01497				
Catagory * Citation of document, with indication, where appropriate, of the relevant passages Re				
	SUNGZOON CHO ET AL: "Evolution of neural network training set through addition of virtual samples" PROCEEDINGS OF 1996 IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION (ICEC'96) (CAT. NO.96TH8114), PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION, NAGOYA, JAPAN, 20-22 MAY 1996, ISBN 0-7803-2902-3, 1996, NEW YORK, NY, USA, IEEE, USA, pages 685-688, XPO02048950 see page 685, left-hand column, line 1 - page 686, right-hand column, line 30; figures 1,2	1,9		
	DATUM M S ET AL: "An artificial neural network for sound localization using binaural cues" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, JULY 1996, ACOUST. SOC. AMERICA THROUGH AIP, USA, vol. 100, no. 1, ISSN 0001-4966, pages 372-383, XP002048951 see abstract see page 372, left-hand column, line 1 - right-hand column, line 31 see page 376, left-hand column, line 26 - page 377, left-hand column, line 29; figures 1,2	1-3,9,10		
A	SLICHER A W R ET AL: "An innovative approach to training neural networks for strategic management of construction firms" DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, CAMBRIDGE, UK, 28-30 AUG. 1995, ISBN 0-948749-36-9, 1995, EDINBURGH, UK, CIVIL-COMP PRESS, UK, pages 87-93, XP002048952 see abstract see page 89, left-hand column, line 5 - page 92, left-hand column, line 8; figure 1	1,9		
	·			

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. .ational Application No PCT/DE 97/01497

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
₩0 9511486 A	27-04-95	US 5444796 A AU 7981794 A CA 2174522 A EP 0724750 A US 5590218 A	22-08-95 08-05-95 27-04-95 07-08-96 31-12-96
US 5359699 A	25-10-94	NONE	·

INTERNATIONALER RECHERCHENBERICHT

Invinationales Aldenzeicher

			101/02 9//0149	<u>'' </u>
A. KLASSI IPK 6	FIZERUNG DES ANMELDUNGSGEGENSTANDES G06F15/80			-
Nach der In	ternationalen Patentklassifikation (IPK) öder noch der nationalen Klas	sifikation und der IPK		
	r Intermationalen Patertiklassifikation (IPR) oder noch der nationelen Klassifikation und der IPK KERCHIERTE GEBIETE Interner Kindesterundstoff (Klassifikationssystem und Klassifikationssymbole) GUGF GUGF der informationalen Recharche konsultians elektronische Deternerik (Name der Datenbank und ext. varwendete Suchbegrifte) WESENTLICH ANGESEHENE UNTERLAGEN WESENTLICH ANG			
Recherchies IPK 6		Forleatzungwon Feld C zu Forleatzungwon Fel		
Recherchler	ne aber nicht zum Mindestprüsstott gehörende Veröffentlächungen, son	west diese unter die rec	herchierten Gebiete Islan	
Während de	er internationalen Recherche konsultierte elektronische Datenbank (N	ame der Dalenbank ur	nd evit. verwendate Suchbegi	ifle)
C. ALS WE	SENTLICH ANGESEHENE UNTERLAGEN			
Kategone*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe	der in Setracht komm	endan Teile 6	etr, Anspruch Nr.
A	siehe Seite 10, Zeile 1 - Zeile 3		1	.,9
A	25.0ktober 1994		1	1,9
	-	·/		
		Siehe Anhang	g Patentiamilie	
* Basonder "A* Veröffle aber n "E* Literas Anme "L* Veröffle scheir ander edi oc ausga "O* Veröffle eine B "P* Veröffle dem b	e Kategorien von angegebenen Veröffentlichungen : rätichung, die den allgemeinen Stand der Technik definiert, nicht als besonders bedeutsam anzusehen ist Dokument, das jedoch erst am oder nach dem internationalen idedatum veröffentlicht worden ist rätichung, die geeignet ist, einen Prioritätsanspruch zweifeihaft er- nen zu lassen, oder durch die das Veröffentlichungsdelum einer en im Recharchenbericht genannten Veröffentlichungsdelum einer en im Recharchenbericht genannten Veröffentlichungsdelum einer ein die aus einem anderen Brund angegeben ist (wie führt) intlichung, die sich auf eine mündliche Öffenbarung, senutzung, eine Ausstellung oder andere Maßnehmen bezieht intlichung, die vor dem internationalen Anmeldedatum, aber nach wenspruchten Prioritätsdatum veröffentlichtworden ist	oder dem Priorität Armetolung nicht i Ertindung zugunch Theoris angegete "X" Varöffertilichung vol kann atlein aufgru- grindsrischer Täll "Y" Veröffertilichung vol kann nicht als auf warden, wann die Veröffe nitlichungsi	adatum veröffentlicht worden vollkliert, eondern nur zum Ve- letiegenden Prinzipe oder de- in tet im beeonderer Bedeutung; die nid dieser Veröffentlichung ni- gkeit beruhend betrachtet we- m beeonderer Bedeutung; die erfinderischer Tätigkeit beruh Veröffentlichung mitteiner od nidseer Kategorie in Veröffentlichen Fachmann nahelier	lat und mit der onständnis des dar ihr zugrundellegendan beanspruchte Erfindung chi ale neu oder auf riden beanspruchte Erfindung send betrachtet er mehreren anderen ung gebracht wird und gend ist.
	Abschlusses der Internationalen Recherche Dezember 1997	Absendedatum da 16/12/1		nberichts
Name und I	Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentiaan 2 NL – 2280 HV Rijawijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3018	Bevollmächtigter (Schenke		

INTERNATIONALER RECHERCHENBERICHT

Invernationales Aktenzeichen
PCT/DE 97/01497

		PCT/DE 9	7//0149/
C.(Fortsetzi	ung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht komm-	enden Teite	Betr. Anspruch Nr.
A	SUNGZOON CHO ET AL: "Evolution of neural network training set through addition of virtual samples" PROCEEDINGS OF 1996 IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION (ICEC'96) (CAT. NO.96TH8114), PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION, NAGOYA, JAPAN, 20-22 MAY 1996, ISBN 0-7803-2902-3, 1996, NEW YORK, NY, USA, IEEE, USA, Seiten 685-688, XP002048950 siehe Seite 685, linke Spalte, Zeile 1 - Seite 686, rechte Spalte, Zeile 30; Abbildungen 1,2		1,9
A	DATUM M S ET AL: "An artificial neural network for sound localization using binaural cues" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, JULY 1996, ACOUST. SOC. AMERICA THROUGH AIP, USA, Bd. 100, Nr. 1, ISSN 0001-4966, Seiten 372-383, XP002048951 siehe Zusammenfassung siehe Seite 372, linke Spalte, Zeile 1 - rechte Spalte, Zeile 31 siehe Seite 376, linke Spalte, Zeile 26 - Seite 377, linke Spalte, Zeile 29; Abbildungen 1,2		1-3,9,10
A v	SLICHER A W R ET AL: "An innovative approach to training neural networks for strategic management of construction firms" DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, CAMBRIDGE, UK, 28-30 AUG. 1995, ISBN 0-948749-36-9, 1995, EDINBURGH, UK, CIVIL-COMP PRESS, UK, Seiten 87-93, XP002048952 siehe Zusammenfassung siehe Seite 89, linke Spalte, Zeile 5 - Seite 92, linke Spalte, Zeile 8; Abbildung 1		1,9

INTERNATIONALER RECHERCHENBERICHT

Angaban zu Veröttentlichungen, die zur selben Patentlamitie gehören

nusemationales Aktenzeichen
PCT/DE 97/01497

tm Recherchenbericht ngeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9511486 A	27-04-95	US 5444796 A AU 7981794 A CA 2174522 A EP 0724750 A US 5590218 A	22-08-95 08-05-95 27-04-95 07-08-96 31-12-96
US 5359699 A	25-10-94	KEINE	

Formblatt PCT/ISA/210 (Anhang Patenthamilie)(Juli 1992)